NATIONAL URBAN LEAGUE

Hearing on H.R. 2622, the Fair and Accurate Credit Transaction Act of 2003
Before the
Committee on Financial Services
United States House of Representatives

TESTIMONY
Concerns about the Transparency of Credit Scores

July 9, 2003

By

William E. Spriggs
Executive Director, National Urban League Institute for Opportunity and Equality

The Urban League is the nation's oldest and largest community-based movement devoted to empowering African Americans to enter the economic and social mainstream.

The Urban League movement was founded in 1910. The National Urban League, headquartered in New York City, spearheads our nonprofit, nonpartisan, community-based movement. The heart of the Urban League movement is our professionally staffed Urban League affiliates in over 100 cities in 34 states and the District of Columbia.

The mission of the Urban League movement is to enable African Americans to secure economic self-reliance, parity and power and civil rights. On behalf of the League, I thank Chairmen Oxley and ranking member Congressman Frank for this opportunity to share the thoughts of the League on this important topic.

The League is encouraged that the proposed H.R. 2622, the Fair Credit Reporting Act of 2003 includes recognition of the need to provide more information to consumers about their credit score. Title V of the proposed Act, amends Section 609(a)(1) of the Fair Credit Reporting Act so individuals can get a summary of their credit score and information on how it was derived so that the score can be improved. We applaud the Committee for that step.

While home ownership is at record levels in the United States, big gaps remain between the homeownership rates of whites, Hispanics and African Americans. The latest data from the Census Bureau ("Census Bureau Reports on Residential Vacancies and

Homeownership," April 24, 2003)[1] shows that as of the first quarter of 2003, 75.0 percent of white non-Hispanic households were homeowners, but that for African Americans the rate was 47.7 percent and for Hispanics it was 46.7 percent. (Homeownership rates also vary by region, with the Northeast and West generally lagging states in the South and Midwest.)

Some of that gap is related to the disparity in home loan rejections faced by African Americans and Hispanics relative to whites. Looking at Home Mortgage Disclosure Act (HMDA) data, researchers Stephen Ross and John Yinger (Syracuse University, "Looking the Other Way: A Critique of the Fair-Lending Enforcement System and a Plan to Fix It")[2] find that the African American to white loan-denial ratio stayed relatively constant between 1995 and 2000 at around 2.0, and the Hispanic to white loan-denial ratio stayed near 1.5.

Beginning in the 1990s credit scores have become more important in deciding which consumers will have access to the primary mortgage market. The pace increased greatly in the latter part of the 1990s, according to work done by Freddie Mac economist John Straka ("A Shift in the Mortgage Landscape: the 1990s Move to Automated Credit Evaluations," *Journal of Housing Research* Vol. *11* (No. 2, 2000)[3] Generally, this development has been greeted as an attempt to reduce discrimination in mortgage lending by making the lending decision objective. But, the Ross and Yinger data suggest that the growing reliance of credit scores and automated underwriting has not changed the loan-denial ratio.

The persistence of the gap in mortgage lending experiences for whites, African Americans and Hispanics, suggest that disparate treatment—that is treating African American and Hispanics differently than whites—may not be important. But, the potential remains for disparate impact—differences in outcomes for African Americans and Hispanics compared to whites because of loan criteria that evaluate equally creditworthy African American and Hispanics less favorably.

How can we avoid the issue of disparate impact of credit scores? If the legislation being considered here goes further, and requires transparency in the score that we also get in the HMDA data. This is not simply an issue of racial fairness that is raised because of the National Urban League's mission. Of course, the issue goes to the heart of the Fair Credit Reporting Act, and that is making credit markets work by making the data as accurate as possible—this benefits both borrowers and lenders, reducing risks to lenders but increasing the share of truly creditworthy individuals who can leverage their assets.

There is a lot of research that the use of credit scoring has created. I will briefly summarize some issues that come under the purview of this Act. They are raised because they highlight the need for transparency in the models, so that researchers may be free to

[1] http://www.census.gov/hhes/www/housing/hvs/q103prss.pdf
[2] http://faculty.maxwell.syr.edu/jyinger/ppa730/looking%20the%20other%20way-%20Final%202.pdf
[3] http://www.fanniemaefoundation.org/programs/jhr/pdf/jhr_1102_straka.pdf

debate the virtues of the models, and policy makers can understand how to balance the needs of accuracy and fairness in the models.

Credit scores are statistical models that use information from credit bureaus to rank the risk of an individual's creditworthiness. As statistical models, they are faced with all the issues of any statistical model, including a desire to be parsimonious yet complete, accurate yet cost-effective. To accomplish these goals, statistical modelers must make judgment calls on which variables to place in the model—since adding variables increase costs.

Keeping the model small means the possibility of "omitted variable" bias. When a researcher leaves out a variable important to the model that is correlated with what the model predicts they will overestimate the importance of the variables in the model. For instance, many credit-scoring models leave out rent paying histories, because these are not reported to credit bureaus. This however, could lead to disparate impacts for groups that are more likely to be renters, like first-time home buyers and those living in regions of high rents like the Northeast, as well as African Americans and Hispanics.

Other omitted variable problems could include employment or health, which may be related to the credit history of an individual, but not be a reliable predictor of future potential. And, as Federal Reserve Board economists Robert Avery, Ralph Bostic and Paul Calem ("Credit Scoring: Statistical Issues and Evidence from Credit-Bureau Files," *Real Estate Economics* Vol. *28* (No. 3, 1996))[4] point out, the models ignore "local economic conditions (such as a regional recession) that may have affected the history of loan repayment in a local area but may be unrelated to future patterns of repayment." They warn that the result could be "different (and unanticipated) levels of credit risk in different parts of the country."

For instance, if local economic conditions are not included in the model, then someone who is unemployed in a depressed county will be treated the same in the model as someone who is employed in a different county who is experiencing the same credit problems. Yet, clearly, if the worker in the depressed county gets a job, or moves to a less depressed area, they may be the better credit risk.

When models are transparent, researchers may evaluate the model, and test the size of the omitted variable bias, and the extent to which the bias creates disparate impacts, either across regions, or racial or income groups. Without that transparency, policy makers, borrowers and lenders are in a fog as to this problem.

Oddly, to remove unintended racial disparate impact from the model, it may be necessary to include the race of the individual in the statistical model. Ross and Yinger point out this is because, if race is excluded from the model, but credit characteristics are different across racial groups enough to be correlated with race, then race can become an omitted variable. For instance, given the huge gap in home ownership rates, mortgage payments themselves are correlated with race. So, some of the measure of the effect of the history

[4] http://www.areuea.org/publications/ree/articles/V28/REE.V28.3.7.PDF

on making mortgage payments the model will be picking up is from the importance of mortgage payments, and another portion will be factors associated with race. The result is that it is necessary to include race in the initial model, and then ignore the race variable in weighting the various factors in the credit-score.

Another problem is how to handle "errors in measurement." Statistical modelers assume they are handling accurate data. Yet, the Consumer Federation of America ("Credit Score Accuracy and Implications for Consumers," December 17, 2002)[5] provides a summary of research done on credit bureau data accuracy, finding a wide range of measures of inaccuracy in the data. Straka discusses the issue of data accuracy, so credit-score modelers are well aware of the issue. But, it is necessary for independent researchers to see how this issue is handled.

A related issue is how the credit-scoring models handle missing data. The various reports cited in the Consumer Federation of America's report show that many errors in reports result from credit card or mortgage lenders who do not file reports. The Treasury Department's Office of the Comptroller of the Currency's Federal Financial Institutions Examination Council felt obliged to issue an advisory letter in January, 2000 because of the drop-off in reports on credit cards and sub-prime loans.[6] How a model handles missing data is crucial, as there are various methods for imputing the missing information, and some can have disparate impacts. For instance, if the model simply drops the missing data, assuming that individuals without mortgage loans reported do not have mortgage loans, and a disproportionate share of sub-prime loans that go unreported are held by African Americans then this will have a racially disparate impact.

Disparate impact can also result because models create their weights using data where some groups are underrepresented. The underrepresented group may have a different set of characteristics that better predicts creditworthiness, or though the model may appear to predict well for both the baseline group used to create the model and the underrepresented group, within each group the predictions can have greater variance. As the Consumer Federation of America report argues, from a lender's perspective, the greater variation may not present a problem. On average, the model will perform equally well for both groups. But, from the consumer's perspective, getting an inaccurate score that can either raise their costs of borrowing, or result in outright denial, there is no counterbalancing weight that average performance gives the lender. So, it is important that researchers be able to determine whether models are similar, not just in average performance, but also that the variance in the models is similar. For instance, on July 8, the Fair Isaacs web page reported that a difference in just 21 points on a credit score, from 699 to 720 could cost a consumer 0.662 percentage points on a mortgage (A quote on a 30-year fixed rate mortgage of 5.332 compared to 5.994 percent).[7]

What can this legislation do? Require the Federal Trade Commission to oversee credit scores, including sunshine provisions to make the inner workings of the models

[5] http://www.consumerfed.org/121702CFA_NCRA_Credit_Score_Report_Final.pdf
[6] http://www.occ.ustreas.gov/ftp/bulletin/2000-3a.txt
[7] http://www.myfico.com/

transparent. The FTC should provide report cards showing: differences in the mean prediction error for subgroups of the population to ensure compliance with all relevant provisions of the Fair Housing Act and the Equal Credit Opportunity Act; detailed notes on how missing values are handled; detailed notes on steps taken to handle errors in measurement in the data; detailed notes on the relative performance of models considered, but rejected, as to their mean prediction error for subgroups of the population with reference to enforcement of the Fair Housing Act and the Equal Credit Opportunity Act, low-to-moderate income, first time homebuyers and rural households, and other relevant characteristics of the models performance. Congress should commission a study of the score models, and further determine, the maximum amount of information about the models that can be made public, with an eye toward further amendments to the Fair Credit Reporting Act to ensure transparency for consumers and regulators.

Credit scoring can add a vital and important link to credit access. But, it must be done carefully, always with an eye toward being as accurate as possible, just as we require from credit bureaus. In the end, America will win by taking full advantage of the new information technology explosion that has transformed the home mortgage industry.